

# A Robust UCB Scheme for Active Learning in Regression from Strategic Crowds

Divya Padmanabhan<sup>1</sup>, Satyanath Bhat<sup>1</sup>, Dinesh Garg<sup>2</sup>, Shirish Shevade<sup>1</sup>, and Y. Narahari<sup>1</sup>

<sup>1</sup> Indian Institute of Science, Bangalore,

<sup>2</sup> IBM India Research Labs

**Abstract.** We study the problem of training an accurate linear regression model by procuring labels from multiple noisy crowd annotators, under a budget constraint. We propose a Bayesian model for linear regression in crowdsourcing and use variational inference for parameter estimation. To minimize the number of labels crowdsourced from the annotators, we adopt an active learning approach. In this specific context, we prove the equivalence of well-studied criteria of active learning like entropy minimization and expected error reduction. Interestingly, we observe that we can decouple the problems of identifying an optimal unlabeled instance and identifying an annotator to label it. We observe a useful connection between the multi-armed bandit framework and the annotator selection in active learning. Due to the nature of the distribution of the rewards on the arms, we use the Robust Upper Confidence Bound (UCB) scheme with truncated empirical mean estimator to solve the annotator selection problem. This yields provable guarantees on the regret. We further apply our model to the scenario where annotators are strategic and design suitable incentives to induce them to put in their best efforts.

## 1 Introduction

Crowdsourcing platforms such as Amazon Mechanical Turk are becoming popular avenues for getting large scale human intelligence tasks executed at a much lower cost. In particular, they have been widely used to procure labels to train learning models. These platforms are characterized by a large pool of diverse yet inexpensive annotators. To leverage these platforms for learning tasks, the following issues need to be addressed: (1) A learning model that encompasses parameter estimation and annotator quality estimation. (2) Identifying the best yet minimal set of instances from the pool of unlabeled data. (3) Determining an optimal subset of annotators to label the instances. (4) Providing suitable incentives to elicit best efforts from the chosen annotators under a budget constraint. We provide an end to end solution to address the above issues for a regression task.

Identifying the best yet minimal set of instances to be labeled is important to minimize the generalization error, as the learner only has limited budget.

This involves selection of those unlabeled instances, the labels of which when fed to the learner, yield maximum performance enhancement of the underlying model. The question of choosing an optimal set of unlabeled examples occupies center stage in the realm of active learning. Past work on active learning in crowdsourcing apply to classification [27, 23] and most of these do not directly apply to regression where the space of labels is unbounded. For instance, the Markov Decision Processes (MDP) based method [23] relies on label space and thereby the state space being finite, which is not the case in regression.

Similar to the instance selection problem, the annotator choice to label an instance also has a bearing on the accuracy of the learnt model. Optimal annotator selection, in the context of classification, has been addressed using multi-armed bandit (MAB) algorithms [1]. Here the annotators are considered as the arms and their qualities as the stochastic rewards. In classification, the quality of the annotators is modeled as a Bernoulli random variable, thereby making it suitable for application of algorithms such as UCB1 [2, 7]. However for regression tasks, the labels provided by the annotators are naturally modeled to have Gaussian noise, the variance of which is a measure of the quality of the annotator. This in turn is a function of the effort put in. Therefore, optimal annotator set selection problem involves identifying annotators with low variance. Though existing work has adopted MAB algorithms for estimating variance [22] and several other applications [30], there is a research gap in its applicability to active learning and regression tasks and in particular where heavy tailed distributions arise as a result of squaring the Gaussian noise. To bridge this gap, we invoke ideas from Robust UCB [8] and set up theoretical guarantees for annotator selection in active learning.

Another non-trivial challenge emerges when we are required to account for the strategic behavior of the human agents. An agent, in the absence of suitable incentives, may not find it beneficial to put in efforts while labeling the data. To induce best efforts from agents, the learner could appropriately incentivize them. In the field of mechanism design, several incentive schemes exist [13, 34]. To the best of our knowledge, such schemes have not been explored in the context of active learning for regression.

**Contributions:** The key contributions of this paper are as follows.

- (1)**Bayesian model for Regression:** In Section 3, we set up a novel Bayesian model for regression using labels from multiple annotators with varying noise levels, which makes the problem challenging. We use variational inference for parameter estimation to overcome intractability issues.
- (2)**Active learning for crowd regression and decoupling instance selection and annotator selection:** In Section 4.1, we focus on various active learning criteria as applicable to the proposed regression model. Interestingly, in our setting, we show that the criteria of *minimizing estimator error* and *minimizing estimator's entropy* are equivalent. These criteria also remarkably enable us to decouple the problems of instance selection and annotator selection.
- (3)**Annotator selection with multi-armed bandits:** In Section 4.2, we de-

scribe the problem of selecting an annotator having least variance. We establish an interesting connection of this problem to the multi-armed bandit problem. In our formulation, we work with the square of the label noise to cast the problem into a variance minimization framework; the square of the noise follows a sub-exponential distribution. We show that standard UCB strategies based on  $\psi$ -UCB [7] are not applicable and we propose the use of robust UCB [8] with truncated empirical mean. We show that the logarithmic regret bound of robust UCB is preserved in this setting as well. Moreover the number of samples discarded is also logarithmic.

(4)**Handling strategic agents:** In Section 5, we consider the case of strategic annotators where the learner needs to induce them to put in their best efforts. For this, we propose the notion of ‘*quality compatibility*’ and introduce a payment scheme that induces agents to put in their best efforts and is also individually rational.

(5)**Experimental validation:** We describe our experimental findings in Section 6. We compare the RMSE and regret of our proposed models with state-of-the-art benchmarks on several real world datasets. Our experiments demonstrate a superior performance.

## 2 Related Work

A rich body of literature exists in the field of active learning for statistical models where labels are provided by a single source [28, 12, 9, 10]. Popular techniques include minimizing the variance or uncertainty of the learner, query by committee schemes [33] and expected gradient length [32] to name a few. In the literature on Optimal Experimental Design in Statistics, the selection of most informative data instances is captured by concepts such as A-optimality, D-optimality, etc. [16, 18]. The idea is to construct confidence regions for the learner and bound these regions. A survey on active learning approaches for various problems is presented in [31].

The works that have looked into active learning for regression are applicable only for a single noisy source, and not to a crowd. In crowdsourcing, several learning models for regression have been proposed, for instance, [25, 26] obtain the maximum likelihood estimate (MLE) and maximum-a-posteriori (MAP) estimate respectively. [17] proposes a scheme to aggregate information from multiple annotators for regression using Gaussian Processes. [4, 24] develop models for classification using crowds. However, these do not employ techniques from active learning. Also, they do not obtain a posterior distribution over the parameters, and hence do not perform probabilistic inference. Of late, there have been a few crowdsourcing classification models employing the active learning paradigm [27, 36, 35, 23, 15]. These include uncertainty-based methods and MDPs. To the best of our knowledge, active learning for regression using the crowds has not been looked at explicitly.

When an annotator is requested to label an instance, and the annotator, being strategic, does not put in the best effort, the learning algorithm could

seriously underperform. So we must incentivize the annotator to induce the best effort. Such studies are not reported in the current literature. [11, 14] propose payment schemes for linear regression for crowds. Both [11, 14] make the assumption that an instance is provided only to a single annotator and also do not look at the active learning paradigm. The idea in our work is to design incentives for active learning in the context of crowdsourced regression which would induce the annotators to put in their best efforts.

In the next section, we explain our model for regression using the crowd, assuming non-strategic annotators.

### 3 Bayesian Linear Regression from a Non-strategic Crowd

Given a data instance  $\mathbf{x} \in \mathbb{R}^d$ , the linear regression model aims at predicting its label  $y$  such that  $y = \mathbf{w}^\top \mathbf{x}$ . Instead of  $\mathbf{x}$ , non-linear functions  $\Phi(\cdot)$  of  $\mathbf{x}$ , can be used. To avoid notational clutter, we work with  $\mathbf{x}$  throughout this paper. The coefficient vector  $\mathbf{w} \in \mathbb{R}^d$  is unknown and training a linear regression model essentially involves finding  $\mathbf{w}$ . Let  $\mathcal{D}$  be the initially procured training dataset and let  $\mathcal{U}$  denote the pool of unlabeled instances. We later (in Section 4.1) select instances from  $\mathcal{U}$  via active learning to enhance our model.

In classical linear regression, the labels are assumed to be provided by a single noisy source. In crowdsourcing, however, there are multiple annotators denoted by the set  $S = \{1, \dots, m\}$ . Each of the annotators provides a label vector which we denote by  $\mathbf{y}_1, \dots, \mathbf{y}_m$ , where  $\mathbf{y}_j \in \mathbb{R}^n$  for  $j = 1, \dots, m$ . Each annotator may or may not provide the label for every instance in the training set. We, therefore, define an indicator matrix  $I \in \{0, 1\}^{n \times m}$ , where  $I_{ij} = 1$  if annotator  $j$  labels instance  $\mathbf{x}_i$ , else  $I_{ij} = 0$ . We denote by  $n_j$ , the number of labels provided by annotator  $j$ . That is,  $n_j = \sum_i I_{ij}$ . We also define a matrix  $\mathbf{X}^j \in \mathbb{R}^{n_j \times d}$  whose rows contain the instances that are labeled by annotator  $j$ . Also, we denote by  $y_{ij}$ , the label provided by annotator  $j$  for  $\mathbf{x}_i$ , which is the same as  $i^{\text{th}}$  element of the label vector  $\mathbf{y}_j$ . The true label of a data instance  $\mathbf{x}_i$  is given by  $\mathbf{w}^\top \mathbf{x}_i$ . Each annotator  $j$  introduces a Gaussian noise in the label he provides. That is,  $y_{ij} \sim \mathcal{N}(\mathbf{w}^\top \mathbf{x}_i, \beta_j^{-1})$  where,  $\beta_j$  is the precision or inverse variance of the distribution followed by  $y_{ij}$ . Intuitively,  $\beta_j$  is directly proportional to the effort put in by annotator  $j$ . We assume that there is always a maximum level of effort that annotator  $j$  can put in and inverse variance corresponding to his best effort is given by  $\beta_j^*$ , which is unknown to the learner as well as other annotators. In general, an annotator may be strategic and may exert a lower effort level  $\beta_j < \beta_j^*$  if appropriate incentives are not provided. In this section, however, we adhere to the assumption that annotators are non-strategic and annotator  $j$  always introduces a precision of  $\beta_j^*$ , thereby setting  $\beta_j = \beta_j^*$ . The parameters of the linear regression model from crowds, therefore, become  $\Theta = \{\mathbf{w}, \beta_1, \dots, \beta_m\}$ . The aim of training a linear regression model is to obtain estimates of  $\Theta$  using the training data  $\mathcal{D}$ . We now describe a Bayesian framework for this.

**Bayesian Model and Variational Inference for Parameter Estimation:**

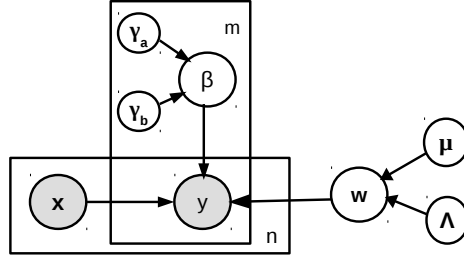


Fig. 1: Plate notation for our model

A Bayesian framework for parameter estimation is well suited for active learning as incremental learning can be done conveniently. Bayesian framework has been developed for estimating the parameters of the linear regression model when labels of training data are supplied by a single noisy source [5]. To the best of our knowledge, the counterpart of such a Bayesian framework in the presence of multiple annotators has not been explicitly explored. We assume a Gaussian prior for  $\mathbf{w}$  with mean  $\mu_0$  and precision matrix or inverse covariance matrix  $\Lambda_0$ . We assume Gamma priors for  $\beta_j$ 's, that is,  $p(\mathbf{w}) \sim \mathcal{N}(\mu_0, \Lambda_0^{-1})$ ,  $p(\beta_j) \sim \mathcal{G}(\gamma_{a0}^j, \gamma_{b0}^j)$  for  $j = 1, \dots, m$ . The plate notation of the Bayesian model described above is provided in Figure 1. The computation of the posterior distributions  $p(\mathbf{w} \mid \mathcal{D})$  and  $p(\beta_j \mid \mathcal{D})$  for  $j = 1, \dots, m$  is not tractable. Therefore, we appeal to variational approximation methods [3]. These methods approximate the posterior distributions using mean field assumptions. We use  $q(\mathbf{w})$  and  $q(\beta_j)$  to represent the mean field variational approximation of  $p(\mathbf{w} \mid \mathcal{D})$  and  $p(\beta_j \mid \mathcal{D})$  respectively. The variational approximation begins by initializing the parameters of the prior distributions,  $\{\mu_0, \Lambda_0\}$  and  $\{\gamma_{a0}^j, \gamma_{b0}^j\}$  for all  $j = 1, \dots, m$ . At each iteration of the algorithm, the parameters of the posterior approximation are updated and the steps are repeated until convergence.

**Lemma 1.** *The variational update rules for the posterior approximations using mean field assumptions are  $q(\mathbf{w}) \sim \mathcal{N}(\mu_n, \Lambda_n^{-1})$  and  $q(\beta_j) \sim \mathcal{G}(\gamma_{an}^j, \gamma_{bn}^j)$  where*

$$\Lambda_n = \left[ \Lambda_0 + \sum_j \mathbb{E}[\beta_j] \sum_{i: I_{ij}=1} \mathbf{x}_i \mathbf{x}_i^\top \right] \quad (1)$$

$$\mu_n = \Lambda_n^{-1} \left[ \Lambda_0 \mu_0 + \sum_j \mathbb{E}[\beta_j] \sum_{i: I_{ij}=1} y_{ij} \mathbf{x}_i \right] \quad (2)$$

$$\gamma_{an}^j = \gamma_{a0}^j + n_j/2 \quad (3)$$

$$\begin{aligned} \gamma_{bn}^j &= \gamma_{b0}^j + \frac{1}{2} \sum_{i: I_{ij}=1} (y_{ij}^2 - y_{ij} \mu_n^\top \mathbf{x}_i) \\ &\quad + \frac{1}{2} \text{Tr} \left( \mathbf{X}^j \mathbf{X}^j \Lambda_n^{-1} \right) + \frac{1}{2} \mu_n^\top \mathbf{X}^j \mathbf{X}^j \mu_n \end{aligned} \quad (4)$$

*Proof.* If  $p$  and  $q$  denote the true and approximate posterior joint distributions of the parameters respectively, we know that,  $\ln p(\mathcal{D}) = \mathcal{L}(q) + KL(q \parallel p)$ , where,

$\mathcal{L}(q) = \int q(\Theta) \ln \left\{ \frac{p(\mathcal{D}, \Theta)}{q(\Theta)} \right\} d\Theta$  and  $KL(q||p) = - \int q(\Theta) \ln \left[ \frac{p(\Theta|\mathcal{D})W}{q(\Theta)} \right] d\Theta$  is the KL divergence between the distributions  $q$  and  $p$ . By the mean field assumption, the joint distribution  $q(\mathbf{w}, \beta_1, \dots, \beta_m)$  factorizes as follows,  $q(\mathbf{w}, \beta_1, \dots, \beta_m) = q(\mathbf{w}) \prod_{j=1}^m q(\beta_j)$ . For simplicity we denote by  $q_{\mathbf{w}}$  the distribution  $q(\mathbf{w})$  and by  $q_{\beta_j}$  the distribution  $q(\beta_j)$ .

$$\begin{aligned} \mathcal{L}(q) &= \int q_{\mathbf{w}} \prod_{i=1}^m q_{\beta_j} \left\{ \ln p(\mathcal{D}, \Theta) - \sum_{i \in \mathbf{w}, \beta_1, \dots, \beta_m} q_i \right\} d\beta_j d\mathbf{w} \\ &\propto \int q_{\mathbf{w}} \left\{ \int \ln p(\mathcal{D}, \Theta) \prod_{i=1}^m q_{\beta_i} d\beta_i \right\} d\mathbf{w} - \int q_{\mathbf{w}} \ln q_{\mathbf{w}} d\mathbf{w} \\ &= \int q_{\mathbf{w}} \ln \tilde{p}(\mathcal{D}, \mathbf{w}) d\mathbf{w} - \int q_{\mathbf{w}} \ln q_{\mathbf{w}} d\mathbf{w} \end{aligned} \quad (5)$$

where,  $\tilde{p}(\mathcal{D}, \mathbf{w}) = E_{\beta} [\ln p(\mathcal{D}, \Theta)] + \text{constant}$  and  $\beta = \{\beta_1, \dots, \beta_m\}$ . In order to minimize  $KL(q||p)$ , we must maximise  $\mathcal{L}(q)$ . Eqn (5) shows that  $\mathcal{L}(q)$  is the negative KL-divergence between  $\tilde{p}(\mathcal{D}, \mathbf{w})$  and  $q_{\mathbf{w}}$ .  $\mathcal{L}(q)$  is maximised when the KL-divergence between  $\tilde{p}(\mathcal{D}, \mathbf{w})$  and  $q_{\mathbf{w}}$  is minimized. Therefore, we must set  $q_{\mathbf{w}} = \tilde{p}(\mathcal{D}, \mathbf{w}) = E_{\beta} [\ln p(\mathcal{D}, \Theta)]$ . By similar calculations, we must set,  $q_{\beta_j} = \tilde{p}(\mathcal{D}, \beta_j) = E_{\mathbf{w}, \beta_{-j}} [\ln p(\mathcal{D}, \Theta)]$ , where  $\beta_{-j} = \{\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_m\}$ .

$$\begin{aligned} \log q(\mathbf{w}) &\propto E_{\beta} [\log p(Y, \mathbf{w}, \beta \mid \mathbf{X}, \delta, \gamma)] \\ &= E_{\beta} [\log p(\mathbf{w}, \beta) + \log p(Y \mid \mathbf{X}, \mathbf{w}, \beta)] \\ &= \log p(\mathbf{w} \mid \delta) + E_{\beta} [\log p(\beta \mid \gamma)] + E_{\beta} [\log p(Y \mid \mathbf{X}, \mathbf{w}, \beta)] \\ &\propto \frac{1}{2\pi} |\Lambda_0| - \frac{1}{2} (\mathbf{w} - \mu_0)^{\top} \Lambda_0 (\mathbf{w} - \mu_0) \\ &\quad + E_{\beta} \left[ \log \prod_{ij} \frac{\beta_j}{2\pi} \exp \left( -\frac{\beta_j (y_{ij} - \mathbf{w}^{\top} x_i)^2}{2} \right) \right] \\ &\propto -\frac{1}{2} (\mathbf{w} - \mu_0)^{\top} \Lambda_0 (\mathbf{w} - \mu_0) \\ &\quad + E_{\beta} \left[ \sum_{ij} \log \frac{\beta_j}{2\pi} - \left( \frac{\beta_j (y_{ij} - \mathbf{w}^{\top} x_i)^2}{2} \right) \right] \\ &= -\frac{1}{2} (\mathbf{w} - \mu_0)^{\top} \Lambda_0 (\mathbf{w} - \mu_0) \\ &\quad + \sum_{ij} E_{\beta_j} \left[ \log \frac{\beta_j}{2\pi} - \left( \frac{\beta_j (y_{ij} - \mathbf{w}^{\top} x_i)^2}{2} \right) \right] \\ &\propto -\frac{1}{2} (\mathbf{w} - \mu_0)^{\top} \Lambda_0 (\mathbf{w} - \mu_0) \\ &\quad + \sum_j E_{\beta_j} \left[ - \left( \frac{\beta_j \sum_i (y_{ij} - \mathbf{w}^{\top} x_i)^2}{2} \right) \right] \end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2}(\mathbf{w}^\top \Lambda_0 \mathbf{w} - 2\mathbf{w}^\top \Lambda_0 \mu_0 + \mu_0^\top \Lambda_0 \mu_0 \\
&\quad + \sum_{ij} \frac{E[\beta_j]}{2}(y_{ij}^2 + x_i^\top \mathbf{w} \mathbf{w}^\top x_i - 2y_{ij} \mathbf{w}^\top x_i) \\
&\propto -\frac{1}{2} \mathbf{w}^\top \left[ \Lambda_0 + \left( \sum_j E[\beta_j] \sum_{i:I_{ij}=1} \mathbf{x}_i \mathbf{x}_i^\top \right) \right] \mathbf{w} \\
&\quad + \mathbf{w}^\top \left[ \Lambda_0 \mu_0 - \frac{1}{2} \sum_j E[\beta_j] \sum_{i:I_{ij}=1} y_{ij} \mathbf{x}_i \right]
\end{aligned}$$

By completing the squares we get the update rules for  $\mathbf{w}$ . The similar steps can be performed to get the variational updates for  $\beta_j$ . Due to constraints on space, we have not included the steps.

The variational updates for  $\mu_n$  and  $\Lambda_n$  defined in Eqns (1) and (2) involve  $\mathbb{E}[\beta_j] = \gamma_{an}^j / \gamma_{bn}^j$ . The updates for  $\gamma_{bn}^j$  given in Eqn (4) involve  $\mu_n$  and  $\Lambda_n$ . This interdependency between the update equations leads to an iterative algorithm.

*Remark 1 (Parameter Estimation).* : Our approach is not tied to the variational inference approximation scheme. For example, MCMC can be used instead.

**Lemma 2. Asymptotic convergence of Bayes estimators:** Let  $\mathbf{w}^*$  be the true underlying value of  $\mathbf{w}$  and the Bayes estimator for  $\mathbf{w}$  under the least squares loss be  $\mu_n$ . Then,  $\lim_{n \rightarrow \infty} \mathbb{E}_{\mathcal{D}}[\mu_n] \rightarrow \mathbf{w}^*$ .

*Proof.* Let  $\mu_n$  and  $\Lambda_n$  be the mean and precision respectively, of the approximate posterior distribution  $q(w)$ , estimated from the training set  $\mathcal{D}$ . Let  $\mathbf{w}^*$  be the realized value of the underlying  $\mathbf{w}$ .

$$\begin{aligned}
E_{\mathcal{D}}[\mu_n] &= \Lambda_n^{-1} (\Lambda_0 \mu_0 + \sum_j E[\beta_j] \sum_{i:I_{ij}=1} \mathbf{x}_i E_{\mathcal{D}}[y_{ij}]) \\
&= \Lambda_n^{-1} (\Lambda_0 \mu_0 + (\Lambda_n - \Lambda_0) \mathbf{w}^*) \\
&= \mathbf{w}^* + \Lambda_n^{-1} \Lambda_0 (\mu_0 - \mathbf{w}^*)
\end{aligned} \tag{6}$$

If the second term in Eqn 6 approaches 0 as  $n \rightarrow \infty$ , the estimate  $\mu_n$  is an asymptotically unbiased estimate for  $\mathbf{w}$ . Using standard linear algebra results, we can prove that the determinant of the precision matrix  $\det(\Lambda_n)$  approaches  $\infty$  with large number of samples, that is,  $\lim_{n \rightarrow \infty} \det(\Lambda_n) \rightarrow \infty$ . Hence the second term in Eqn 6 approaches zero. Therefore  $\lim_{n \rightarrow \infty} E_{\mathcal{D}}[\mu_n] \rightarrow \mathbf{w}^*$ .

Lemma 2 is a desirable property of the estimators, and in general holds true for Bayes estimators.

**Inference:** We now describe an inference scheme to make prediction about the label of a test data instance. We denote by  $\hat{y}_{\text{test}}$  the predicted label for the test instance  $\mathbf{x}_{\text{test}}$ . From the Bayesian framework of parameter estimation, The posterior predictive distribution for  $\hat{y}_{\text{test}}$  turns out to be as follows:  $p(\hat{y}_{\text{test}} | \mathbf{x}_{\text{test}}, \mathcal{D}) \sim \mathcal{N}(\mathbf{x}_{\text{test}}^\top \mu_n, \mathbf{x}_{\text{test}}^\top \Lambda_n^{-1} \mathbf{x}_{\text{test}})$ . This follows from standard results in [5]. We can use this distribution later in scenarios like active learning.

## 4 Active Learning for Linear Regression from the Crowd

We now discuss various active learning [31] strategies in our framework. Let  $\mathcal{U}$  be the set of unlabeled instances. The goal is to identify an instance, say  $\mathbf{x}_k \in \mathcal{U}$ , for which seeking a label and retraining the model with this additional training example will improve the model in terms of the generalization error. In the crowdsourcing context, since multiple annotators are involved, we also need to identify the annotator  $t$  from whom we should obtain the label for  $\mathbf{x}_k$ . The active learning criterion, thus, involves finding a pair  $(k, t)$  so that retraining with the new labeled set  $\mathcal{D} \cup \{(\mathbf{x}_k, y_{kt})\}$  would provide maximum improvement in the model.

### 4.1 Instance Selection

To our crowdsourcing model, we now apply two criteria well-studied in active learning from a single source. We also show that all these seemingly different criteria embody the same logic.

**Minimizing Estimator Error** Minimizing estimator error is a natural criterion for active learning [29]. The error in the estimator  $\mu_{n+1}$ , if we choose a pair  $(k, t)$ , is given by,  $\text{Err}(\mu_{n+1}) = \mathbb{E}_{y_{kt}} [\mu_{n+1}] - \mathbf{w}$ . The error in the estimator  $\mu_n$ , before including the instance  $(\mathbf{x}_k, y_{kt})$  in the training set is,  $\text{Err}(\mu_n) = \mu_n - \mathbf{w}$ .

**Lemma 3.** *The relation between errors in  $\mu_{n+1}$  and  $\mu_n$  is given by,*

$$\|\text{Err}(\mu_n)\| / (1 + \beta_t \mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k) \leq \|\text{Err}(\mu_{n+1})\| \leq \|\text{Err}(\mu_n)\| \quad (7)$$

*Proof.* We first compute  $E_{y_{kt}} [\mu_{n+1}]$ .

$$\begin{aligned} E_{y_{kt}} [\Lambda_{n+1} \mu_{n+1}] &= \Lambda_{n+1} E_{y_{kt}} [\mu_{n+1}] \\ &= \Lambda_n \mu_n + \mathbf{x}_k (\mathbf{x}_k^\top \mathbf{w}) \beta_t \end{aligned} \quad (8)$$

Making necessary substitutions and rearranging the terms,

$$E_{y_{kt}} [\mu_{n+1}] - \mu_n = -\Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \text{Err}(\mu_{n+1}) \beta_t$$

Again rearranging the terms and subtracting  $\mathbf{w}$  from both the sides yields,  $\text{Err}(\mu_{n+1}) = (I + \Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \beta_t)^{-1} \text{Err}(\mu_n)$ . We now bound  $\text{Err}(\mu_{n+1})$ , in terms of the old error,  $\text{Err}(\mu_n)$  as follows:  $\|\text{Err}(\mu_{n+1})\| \leq \|(I + \Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \beta_t)^{-1}\| \|\text{Err}(\mu_n)\|$  where,  $\|(I + \Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \beta_t)^{-1}\|$  is the spectral norm of the matrix  $(I + \Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \beta_t)^{-1}$ . Since  $\Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top$  is a rank one matrix, the matrix  $I + \Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \beta_t$  has  $d-1$  eigenvalues equal to 1 and one eigenvalue equal to  $1 + \beta_t \mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k$ . Note,  $\mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k > 0$  since  $\Lambda_n^{-1}$  is a positive definite matrix. Therefore, spectral norm of the matrix  $(I + \Lambda_n^{-1} \mathbf{x}_k \mathbf{x}_k^\top \beta_t)^{-1}$  is 1 and its minimum eigenvalue is  $1/(1 + \beta_t \mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k)$  and we arrive at the error bound.

From Theorem 3, it is clear that to reduce the value of the lower bound, we must pick a pair  $(k, t)$  for which the score  $\beta_t \mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k$  is maximum.



**Minimizing Estimator's Entropy** This is another natural criterion for active learning which suggests that the entropy of the estimator after adding an example should decrease [20, 21]. Formally, let  $H(\mathbf{w} \mid \mathcal{D})$  and  $H(\mathbf{w} \mid \mathcal{D}')$  denote the entropies of the estimator before and after adding an example, respectively, where we have  $\mathcal{D}' = \mathcal{D} \cup \{(\mathbf{x}_k, y_{kt})\}$ . Again, let us assume  $\beta_j$ 's are known for the time being. The entropy of the distribution before adding an example satisfies:  $H(\mathbf{w} \mid \mathcal{D}) \propto \det(\Lambda_n^{-1})$ . After adding the example, entropy function behaves as follows.  $H(\mathbf{w} \mid \mathcal{D}') \propto \det(\Lambda_{n+1}^{-1})$ , where

$$\det(\Lambda_{n+1}^{-1}) = \det(\Lambda_n^{-1}) / (1 + \beta_t \mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k) \quad (9)$$

From (9), we would like to choose an instance  $\mathbf{x}_k$  and an annotator  $t$  that jointly maximize  $\beta_t \mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k$  so that  $\det(\Lambda_{n+1}^{-1})$  as well as estimator's entropy are minimized. Recall, the same selection strategy was obtained while using the *minimize estimator error* criterion. Let  $\lambda^* = \lambda_{\max}(\Lambda_n^{-1})$  and  $\lambda_* = \lambda_{\min}(\Lambda_n^{-1})$ . We can further bound the estimator precision as follows.

$$1/(1 + \beta_t \lambda^* \|\mathbf{x}_k\|^2) \leq \det(\Lambda_{n+1}^{-1}) / \det(\Lambda_n^{-1}) \leq 1/(1 + \beta_t \lambda_* \|\mathbf{x}_k\|^2)$$

We observe that the selection of the best instance  $\mathbf{x}_k$  and the best annotator  $S_t$  can be decoupled. That is, we can first select an instance  $\mathbf{x}_k$  for which  $\mathbf{x}_k^\top \Lambda_n^{-1} \mathbf{x}_k$  is maximum and independently select an annotator for whom  $\beta_t$  is maximum. But this scheme of annotator selection may lead to starvation of best annotators if the annotators have not been explored sufficiently. Hence we only use this strategy for selecting an instance and not for selecting the annotator.

## 4.2 Selection of an Annotator

Having chosen the instance  $\mathbf{x}_k$ , next the learner must decide which annotator should label it. Consider any arbitrary sequential selection algorithm  $A$  for the annotators. If the variance of the annotators' labels were known upfront, the best strategy would be to always select the annotator introducing the minimum variance  $1/\beta^* = \min_{1 \leq j \leq m} 1/\beta_j$ . The variances of the annotators' labels are unknown and hence a sequential selection algorithm  $A$  incurs a regret defined by  $\text{Regret-Seq}(A)$  below. We denote the sub-optimality of annotator  $j$  by  $\Delta_j = (1/\beta_j) - (1/\beta^*)$ .

**Definition 1. *Regret-Seq(A, t):*** If  $T_j(t)$  is the number of times annotator  $j$  is selected in  $t$  runs of  $A$ , the expected regret of  $A$  in  $t$  runs, with respect to the choice of annotator, is computed as,  $\text{Regret-Seq}(A, t) = \sum_{j=1}^m \Delta_j \mathbb{E}[T_j(t)]$ .

The problem is to formally establish an annotator selection strategy which yields a regret as low as possible. The main challenge is that the annotators' noise level is unknown and must be estimated simultaneously while also deciding on the selection strategy. We observe the connections of this problem to the multi-armed bandit (MAB) problem. In MAB problems, there are  $m$  arms each producing rewards from fixed distributions  $P_1, \dots, P_m$  with unknown means  $\gamma_1, \dots, \gamma_m$ .

The goal is to maximise the overall reward and for this, at every time-step a decision has to be made as to which arm must be pulled. We denote the sub-optimality of arm  $i$  by  $\Delta_i^{\text{MAB}} = \gamma_* - \gamma_i$ , where  $\gamma_* = \max_{1 \leq i \leq m} \gamma_i$ .

**Definition 2. *Regret-MAB( $M, t$ ):*** If  $T_i(t)$  is the number of times arm  $i$  is selected in  $t$  runs of any MAB algorithm  $M$ , the expected regret of  $M$  in  $t$  runs, *Regret-MAB( $M, t$ )*, is computed as,  $\text{Regret-MAB}(M, t) = \sum_{i=1}^m \Delta_i^{\text{MAB}} \mathbb{E}[T_i(t)]$ .

We now show that the active learning problem in crowdsourcing regression tasks can be mapped to the MAB problem. We know that,  $\mathbb{E}[(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2] = 1/\beta_j$ . Since we are interested in the annotator introducing the minimum variance, we could work with a MAB framework where the rewards of the arms (annotators in our case) are drawn from the distribution of  $-(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$ . This idea was used in [22] in the context of sequential selection from a pool of Monte Carlo estimators. If the selection strategy  $A$  appeals to any MAB algorithm  $M$  defined on the distributions  $-(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$ ,  $\text{Regret-MAB}(M, t)$  will be the same as  $\text{Regret-Seq}(A, t)$ , as proved by [22]. This implies that for the selection strategy, we could work with any standard MAB algorithm such as UCB on the distribution of  $-(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$  and  $\text{Regret-Seq}(A, t)$  would be the same as  $\text{Regret-MAB}(M, t)$ , for an appropriately formulated MAB algorithm  $M$ .

**UCB Algorithm on  $-(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$**  As mentioned, we can work with MAB algorithms on  $-(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$  for which we look at the widely used UCB family of MAB algorithms. The UCB algorithm is an index based scheme which, at time instant  $t$  selects an arm  $i$  that has the maximum value of sum of the estimated mean ( $\hat{\gamma}_i$ ) and a carefully designed confidence interval  $c_{i,t}$  to provide desired guarantees. To design the UCB confidence interval  $c_{i,t}$ , a fairly general class of algorithms called  $\psi$ -UCB [7] can be used. The procedure for applying  $\psi$ -UCB for a random variable  $G$  with some arbitrary distribution, involves choosing a convex function  $\psi_G(\lambda)$ , such that,  $\ln \mathbb{E}[\exp(|\lambda(G - \mathbb{E}[G])|)] \leq \psi_G(\lambda)$  for all  $\lambda \geq 0$ . Further, an application of Chernoff bounds gives the confidence interval. In particular when  $G$  satisfies the sub-Gaussian property, the choice of  $\psi_G(\lambda)$  is easy. In our setting, we will see that  $\psi$ -UCB is inapplicable.

**Lemma 4. *Inapplicability of  $\psi$ -UCB:*** Let the distribution of random variables  $G_j$  follow a zero-mean normal distribution for  $j = 1, \dots, m$ . The distribution of  $-G_j^2$  is sub-exponential which is a heavy-tailed distribution. For an MAB framework where the rewards of the arms are sampled from  $-G_j^2$ ,  $\psi$ -UCB is not applicable.

*Proof.* A variable  $G$  is sub-exponential if  $\mathbb{E}[\exp(\lambda G)] \leq 1/(1 - \lambda/a)$  for  $0 < \lambda < a$ . We now prove that the random variable  $G^2$ , where  $G \sim \mathcal{N}(0, \sigma^2)$  is

sub-exponential.

$$\mathbb{E}[\exp(\lambda G^2)] = 1/(\sigma\sqrt{2\pi}) \int_{-\infty}^{\infty} \exp(z^2(\lambda - (1/2\sigma^2)))dz \quad (10)$$

$$= 1/(\sigma\sqrt{2\pi}) \int_{-\infty}^{\infty} \exp(-z^2((1 - 2\lambda\sigma^2)/2\sigma^2))dz \quad (11)$$

$$= 1/(\sigma\sqrt{2\pi}) \int_{-\infty}^{\infty} \exp(-z^2/(2\sigma^2/(1 - 2\lambda\sigma^2)))dz \quad (12)$$

$$= 1/\sqrt{1 - 2\lambda\sigma^2} \quad (13)$$

$$< 1/(1 - 2\lambda\sigma^2) \text{ for } 0 \leq \lambda < 1/2\sigma^2 \quad (14)$$

Setting  $a = 1/\sigma^2$  shows that  $G$  is sub-exponential. A random variable  $-G^2$  is sub-exponential iff  $G^2$  is sub-exponential. Therefore  $-G^2$  is sub-exponential.

Let  $G_j \sim \mathcal{N}(0, \sigma_j^2)$ . We now compute the functions,  $\mathbb{E}[\exp(\lambda(-G_j^2 + \mathbb{E}[G_j^2]))]$  and  $\mathbb{E}[\exp(\lambda(\mathbb{E}[-G_j^2] + (G_j^2)))]$ .  $\mathbb{E}[G_j^2] = \sigma_j^2$ .

$$\begin{aligned} \mathbb{E}[\exp(\lambda(-G_j^2 + \mathbb{E}[G_j^2]))] &= \mathbb{E}[\exp(\lambda(-G_j^2 + \sigma_j^2))] \\ &= \frac{\exp(\lambda\sigma_j^2)}{\sigma_j\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-\lambda x^2) \exp(\frac{-x^2}{2\sigma_j^2}) dx \\ &= \exp(\lambda\sigma_j^2)/(\sigma_j\sqrt{2\pi}) \int_{-\infty}^{\infty} \exp(-x^2/2(\sigma_j^2/(1 + 2\lambda\sigma_j^2))) dx \\ &= \exp(\lambda\sigma_j^2)/\sqrt{1 + 2\lambda\sigma_j^2} \end{aligned}$$

Similar calculations also yield,

$$\mathbb{E}[\exp(\lambda(G_j^2 - \mathbb{E}[G_j^2]))] = \exp(-\lambda\sigma_j^2)/\sqrt{1 - 2\lambda\sigma_j^2}$$

In order to apply  $\psi$ -UCB for the MAB framework where the rewards of the arms are sampled from  $-G_j^2$ , we need to compute a function  $\psi(\lambda)$  such that for all  $\lambda \geq 0$ ,  $\ln \mathbb{E}[\exp(\lambda(G_j^2 - \mathbb{E}[G_j^2]))] \leq \psi(\lambda)$  and  $\ln \mathbb{E}[\exp(\lambda(-G_j^2 + \mathbb{E}[G_j^2]))] \leq \psi(\lambda)$ .  $\mathbb{E}[\exp(\lambda(G_j^2 - \mathbb{E}[G_j^2]))]$  is not even defined for  $\lambda \geq 1/(2\sigma_j^2)$  and hence the function  $\psi(\lambda)$  cannot be computed. Therefore  $\psi$ -UCB cannot be applied to this framework.

In our setting,  $y_{kj} - \mathbf{w}^\top \mathbf{x}_k$  follows a normal distribution and  $-(y_{kj} - \mathbf{w}^\top \mathbf{x}_k)^2$  has a sub-exponential distribution which is heavy tailed. Therefore from Lemma 4, an upper confidence interval cannot be obtained using  $\psi$ -UCB.

**Robust-UCB with Truncated Empirical Mean** To devise upper confidence intervals for heavy tailed distributions, Robust UCB [8] prescribes working with ‘robust’ estimators such as a truncated empirical mean, where samples that lie beyond a carefully chosen range are discarded. The necessary condition to be satisfied while applying Robust UCB is that the reward distribution of the arms

should have moments of order  $1 + \epsilon$  for some  $\epsilon \in (0, 1]$ . Since the distribution of  $-(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$  has finite variance, Robust UCB with the truncated empirical mean can be used by setting  $\epsilon = 1$ . At round  $t$ , the truncated empirical mean of the samples, the absolute value of which do not exceed  $\sqrt{ut/\log \delta^{-1}}$ , is computed as,

$$\hat{\mu}_t^j = \frac{1}{n_j^c} \sum_{i: I_{ij}=1} \xi_{ij} \mathbb{1}(|\xi_{ij}| \leq \sqrt{ut/\log \delta^{-1}}) \quad (15)$$

where  $\xi_{ij} = -(y_{ij} - \mu_w^\top \mathbf{x}_i)^2$  and  $\mu_w$  is the estimator of  $\mathbf{w}$  obtained from the variational inference algorithm. In Eqn 15,  $n_j^c$  is the number of samples that are actually considered,  $\delta$  is the desired confidence on the deviation of  $\hat{\mu}_t^j$  from  $1/\beta_j$  for all  $j$ ,  $u$  is an upper bound on  $\xi_{ij}^{1+\epsilon}$ . From Lemma 2  $\mu_w$  is an unbiased estimate for  $\mathbf{w}$  and hence we use  $\mu_w$  instead of  $\mathbf{w}$ . The parameter  $\delta$  can be tuned appropriately to get tight bounds on the regret. We now describe the algorithm.

---

**Input:** No. of annotators  $m$ , Unlabeled set  $\mathcal{U}$ , Labeled set  $\mathcal{D}$ ,  $n_j$ ,  $n_j^c$ , for  $j = 1, \dots, m$   
Set  $\mu_w, \Lambda_w$  using variational inference procedure described earlier;  $t := 0$  ;  
Set  $\hat{\mu}_t^j$  for the annotators using Eqn (15);  
**while** ( the learner has budget or the model has not attained the desired RMSE ) **do**  
    – Choose an instance  $\mathbf{x}_k = \arg \max_{\mathbf{x} \in \mathcal{U}} \mathbf{x}^\top \Lambda_w^{-1} \mathbf{x}$  ;  
    – Get a label  $y_{kj^*}$  from an annotator  $j^*$  such that  $j^* \in \arg \max_{1 \leq j \leq m} \hat{\mu}_t^j + \sqrt{32u(\log t)/n_j}$  ;  
    –  $t := t + 1$  ;  $n_{j^*} := n_{j^*} + 1$  ;  $\mathcal{D} := \mathcal{D} \cup \{(\mathbf{x}_k, y_{kj^*})\}$  ;  
    – Run variational inference procedure described earlier and update  $\mu_w$  ;  
    – If  $(y_{kj^*} - \mu_w^\top \mathbf{x}_i)^2 < \sqrt{ut/\log \delta^{-1}}$   
        •  $n_{j^*}^c := n_{j^*}^c + 1$  ;  
        • Update  $\hat{\mu}_t^{j^*}$  using Eqn (15);  
**end**

---

**Algorithm 1:** Robust UCB for selecting the annotators

---

**Theorem 1.**  $\text{Regret-Seq}(\text{Algo 1}, T) \leq \sum_{i: \Delta_i > 0} \frac{32u \log T}{\Delta_i} + 5\Delta_i$ .

*Proof.* We first prove that, with probability at least  $1 - \delta$ ,

$$\hat{\mu}_t^j \leq (-1/\beta_j) + 4\sqrt{u \log \delta^{-1}/n_j} \quad (16)$$

Let  $C_t = \sqrt{ut/\log \delta^{-1}}$ . Let the random variable  $\xi = -(y_{kj} - (\mathbf{w}^\top \mathbf{x}_k))^2$ . As mentioned earlier  $\xi^{1+\epsilon} = \xi^2 < u$ . Note that

$$\mathbb{E}[\xi^2 \mathbb{1}_{\xi \leq C_t}] = \mathbb{E}[|\xi^2| \mathbb{1}_{\xi \leq C_t}] \leq u \quad (17)$$

$$\begin{aligned}
\mathbb{E}[\xi \mathbb{1}_{|\xi| > C_t}] &\leq \mathbb{E}[|\xi|^2]^{1/2} \mathbb{E}[\mathbb{1}_{\xi \geq C_t}]^{1/2} \leq \sqrt{u} (P\{\xi \geq C_t\})^{1/2} \\
&\leq \sqrt{u} (\mathbb{E}[\xi^2]/C_t^2)^{1/2} = u/C_t
\end{aligned} \tag{18}$$

Equation (18) arises due to Holder's inequality. Further,

$$\begin{aligned}
\mathbb{E}[\xi] - \frac{1}{n_j} \sum_{t=1}^{n_j} \xi_t \mathbb{1}_{[\xi_t \leq C_t]} &= \frac{1}{n_j} \sum_{t=1}^{n_j} (\mathbb{E}[\xi] - \mathbb{E}[\xi \mathbb{1}_{|\xi| \leq C_t}]) \\
&\quad + \frac{1}{n_j} (\mathbb{E}[\xi \mathbb{1}_{|\xi| \leq C_t}] - \xi_t \mathbb{1}_{[\xi_t \leq C_t]}) \\
&= \frac{1}{n_j} \sum_{t=1}^{n_j} \mathbb{E}[\xi \mathbb{1}_{|\xi| > C_t}] + \frac{1}{n_j} (\mathbb{E}[\xi \mathbb{1}_{|\xi| \leq C_t}] - \xi_t \mathbb{1}_{[\xi_t \leq C_t]}) \\
&\leq \frac{u}{C_t} + \sqrt{\frac{2u \log \delta^{-1}}{n_j}} + \frac{2C_n \log \delta^{-1}}{3n_j}
\end{aligned} \tag{19}$$

The first term in Eqn (19) arises as a consequence of Eqn (18) and the remaining terms arise as a result of Bernstein's inequality with some simplification. Further algebraic simplification of Eqn (19) gives us Eqn (16).

For a MAB algorithm  $A$  using  $\hat{\mu}_t^j$  as an estimator for  $-1/\beta_j$ , the regret satisfies the following bound when  $\delta = T^{-2}$ , where  $T$  is the total time horizon of plays of the MAB algorithm.

$$\text{Regret-MAB}(A, T) \leq \sum_{i: \Delta_i > 0} \frac{32u \log T}{\Delta_i} + 5\Delta_i. \tag{20}$$

Proof of Eqn (20) involves bounding the number of trials where a sub-optimal arm is pulled, similar to the technique in [2, 8]. A pull of a sub-optimal arm indicates one of the following three events occur: (1) The mean corresponding to the best arm is underestimated (2) the mean corresponding to a sub-optimal arm is over-estimated (3) the mean corresponding to the sub-optimal arm is close to that of the optimal arm. Next we bound each of the three events and use union bound to get the final result. Eqn (16) is used to get bounds for events (1) and (2).  $\text{Regret-Seq}(\text{Algo 1}, T) = \text{Regret-MAB}(\text{Robust-UCB}, T)$  from [22].

**Theorem 2.** *The expected number of samples discarded by the Robust UCB algorithm in  $t$  trials of the algorithm,  $\mathbb{E}[W(t)] \leq 4(\log t)^2$ .*

*Proof.* As per the robust UCB algorithm, at the  $t^{\text{th}}$  time instant, the probability of the random variable  $\xi = (y_{kj} - \mathbf{w}^\top \mathbf{x}_k)^2$  exceeding  $(ut/(4 \log t))^{1/(1+\epsilon)}$ ,

$$\begin{aligned}
P(\xi > (ut/(4 \log t))^{1/(1+\epsilon)}) &= P(\xi^{1+\epsilon} > ut/(4 \log t)) \\
&\leq \frac{\mathbb{E}[\xi]^{1+\epsilon} 4 \log t}{ut} \quad (\text{by Markov inequality}) \\
&\leq 4 \log t / t
\end{aligned}$$

The number of samples discarded upto a time  $n$  is

$$\begin{aligned}\mathbb{E}[W(n)] &= \sum_{t=1}^n \mathbb{E}[\mathbb{1}[Z_t > (ut/(4 \log t))^{\epsilon/(1+\epsilon)}]] \\ &= \sum_{t=1}^n 4 \log t/t \leq 4(\log n)^2 \quad \square\end{aligned}$$

## 5 The Case of Strategic Annotators

Till now, we have inherently assumed that annotators are non-strategic. Now we look at the scenario where an annotator who has been allocated an instance is strategic about how much effort to put in. For this, we assume that, for each annotator  $j$ , the precision  $\beta_j$  introduced while labeling an instance is proportional to the effort put in by annotator  $j$ . We now refer to the effort as  $\beta_j$  for simplicity. It is best for the learning algorithm when the annotator  $j$  puts in as much effort (high  $\beta_j$ ) as possible thereby reducing the variance in the labeled data. A given level of effort incurs a cost to the annotator  $c_j(\beta_j)$ . We assume that  $c_j(\cdot)$  is a non-negative strictly increasing function of  $\beta_j$  with  $c_j(0) = 0$ . The exact form of  $c_j(\cdot)$  is unknown to the learner. From the annotator's point of view, a high value of effort  $\beta_j$  might incur a higher cost and thus the annotator might not be motivated to put in higher effort.

In order to take into account the strategic play of the human annotators, we appeal to mechanism design techniques. Mechanism design comprises allocation and payment rules. The mechanism is to be designed to meet at least the following objectives.

**Definition 3.** *Individual Rationality (IR): A mechanism is IR if the expected utility of every participating agent is non-negative.*

**Definition 4.** *Quality Compatibility: We say a mechanism is 'quality compatible' at level  $\underline{\beta}$  if it induces every participating agent to operate under precision  $\beta \geq \underline{\beta}$ .*

We now present a mechanism design solution which meets the above design goals.

**Proposed Mechanism:** (1) We use Algorithm 1 as the allocation rule. (2) The payment rule for annotator  $j$  when his estimated precision is  $\hat{\beta}_j$  is,

$$P(\hat{\beta}_j) = B \min \left\{ 1, \max \left\{ 0, \left( \frac{\hat{\beta}_j - \underline{\beta}}{\bar{\beta} - \underline{\beta}} \right) \right\} \right\} \quad (21)$$

We assume that the learner has a finite budget  $B$  per example. Also the annotators are expected to have precisions in the range  $[\underline{\beta}, \bar{\beta}]$ . An effort level in this expected range fetches a corresponding proportional payment to the annotator. If an annotator puts in an effort less than  $\underline{\beta}$ , he does not receive any payment. An effort level higher than  $\bar{\beta}$  fetches an annotator a maximum payment of  $B$ , due to the limitation on the willingness of the learner.

**Annotator's optimization problem:** The utility of the annotator when operating at the effort level  $\beta_j$  is  $U(\beta_j) = P(\beta_j) - c_j(\beta_j)$ . The optimal effort for the annotator,  $\beta_j^* = \operatorname{argmax}_{\beta_j} U(\beta_j)$ .

**Theorem 3.** *The proposed mechanism is IR and quality compatible.*

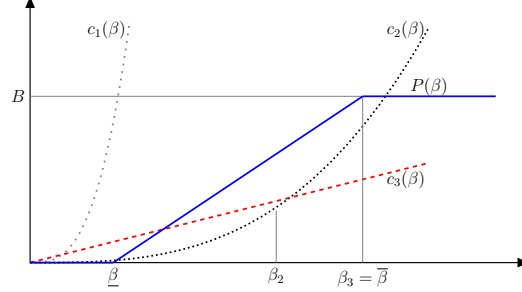


Fig. 2: Realizations of  $c_j(\beta_j)$  in relation to the payment rule  $P(\beta_j)$

*Proof.* The payment scheme is individually rational as annotators participate only when  $P(\beta_j) > c_j(\beta_j)$  in that case, they obtain a positive utility. The utility is therefore non-negative and hence the mechanism is IR.

In order to prove that the payment scheme is quality compatible, we consider the three possible realizations of  $c_j(\beta_j)$  in relation to the payment rule  $P(\beta_j)$ .

1. There exists no  $\beta_j$  for which  $P(\beta_j) > c_j(\beta_j)$ . In this scenario, an annotator will choose to not participate, as there is clearly no benefit from participation. The cost function  $c_1(\beta)$  in Figure 2 captures this.
2. There exists some  $\beta_j$  such that  $\underline{\beta} < \beta_j \leq \bar{\beta}$ , for which  $P(\beta_j) > c_j(\beta_j)$  and the maximum utility is attained at  $\beta_{j^*}$ , such that,  $\underline{\beta} < \beta_{j^*} < \bar{\beta}$ . The cost function  $c_2(\beta)$  in Figure 2 demonstrates this scenario where an effort  $\beta_2 > \underline{\beta}$  maximizes his utility.
3. There exists  $\beta_j$  such that  $\underline{\beta} < \beta_j \leq \bar{\beta}$ , for which  $P(\beta_j) > c_j(\beta_j)$  and the maximum utility is attained at  $\beta_{j^*} = \bar{\beta}$ . The cost function  $c_3(\beta)$  in Figure 2 demonstrates this.  $\square$

## 6 Experimental Results

We conducted experiments on three real world datasets from the UCI repository [19] - Housing, Redwine and Whitewine, the details of which are provided in Table 1a. To simulate the annotators, we added zero-mean Gaussian noise to the output variables.  $1/\sqrt{\beta_j}$  values of the annotators were randomly chosen from two sets of intervals  $U1 = [0.1, 1]$  and  $U2 = [1, 2]$ . Annotators with  $1/\sqrt{\beta_j}$  chosen from interval  $U1$  are clearly better than those chosen from  $U2$ .

### 6.1 Data Preprocessing

We worked with a transformation  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$  of the original data matrix  $\mathbf{X}$ . For the Housing and the Whitewine datasets, we worked with the following

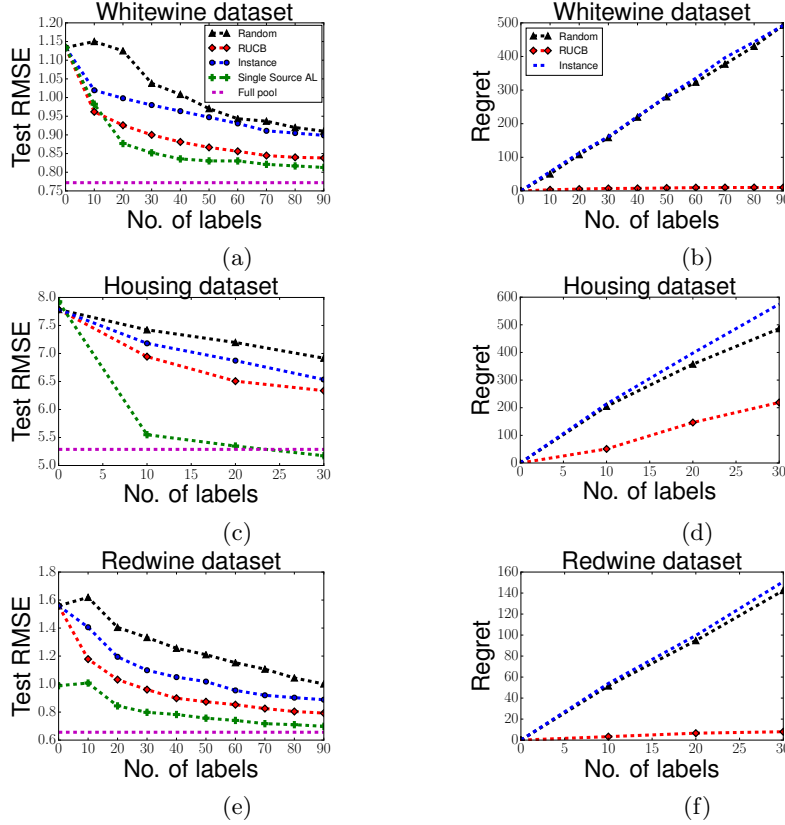


Fig. 3: Active learning results on various datasets. The legends for the figures in each row are provided on the corresponding figure in the first row

non-linear transformation  $\Phi_b(\mathbf{x}; R_b, s) = 1/(1 + \exp(-\|\mathbf{x} - R_b\|/s))$ ,  $b = 1 \rightarrow d$  whereas, for the Redwine dataset, the original data matrix  $\mathbf{X}$  was used. The value of  $s$  was fixed using cross-validation. The parameters  $R_b$  for  $b = 1 \rightarrow d$ , were set as the  $k$ -means cluster representatives of the dataset. All the features were normalized.

## 6.2 Performance of Bayesian Parametric Model

We compared our Bayesian parameter estimation algorithm (without active learning) with MLE [25] and Gaussian Process based method [17]. From the complete dataset  $\mathcal{C}$ , a random 30% of data was used as test dataset  $\mathcal{T}$ . We refer to the set  $\mathcal{C} \setminus \mathcal{T}$  as the full pool of training instances  $\mathcal{F}$ . 50 annotators were used, out of which, for 40 of them, the parameter  $1/\sqrt{\beta_j}$  was chosen from  $U1$ , and for the remaining, from  $U2$ . The parameters of the Bayesian model described earlier were learnt using the full pool  $\mathcal{F}$  labeled by the 50 annotators, as the training data. The experiments were repeated with 10 different splits of the data. We report the Average Root Mean Square Error (RMSE)



scores on the test set  $\mathcal{T}$ . The RMSE for the test dataset containing  $N_{\text{test}}$  instances with true output vector  $\mathbf{z}$  and predicted output vector  $\hat{\mathbf{y}}$ , is calculated as,  $\text{RMSE}(\hat{\mathbf{y}}; \mathbf{z}) = \sqrt{\sum_{i=1}^{N_{\text{test}}} (\hat{y}_i - z_i)^2 / N_{\text{test}}}$ . Our results are provided in Table 1b. Our method consistently outperforms Groot’s method and compares well with MLE.

*Remark 2.* With increasing size of the dataset, the performance of our model approaches MLE (as demonstrated in Table 1b, Whitewine dataset). This is consistent with the result that with increased size of training data set, Bayesian estimates perform similar to MLE [6]. It further shows the efficacy of our learning scheme explained in Section 3. The additional advantage that our model offers is the suitability to further apply active learning methods, which is not offered by other learning schemes like MLE[25] and Groot et al [26].

Dataset	Size	$d$	$\Phi$
Housing	506	12	Nonlinear
Redwine	1599	11	Linear
Whitewine	4898	11	Nonlinear

(a) Details of datasets

Dataset	Our method	MLE	Groot et al.
Housing	<b>4.7209</b>	4.93834	5.998169
Redwine	<b>0.51490</b>	0.65868	0.67354
Whitewine	<b>0.75740</b>	0.75748	1.235

(b) Average test RMSE values when the whole dataset is used (without active learning). ‘Our method’ refers to the variational inference based learning scheme explained in Section 3

Table 1: Details of datasets and performance of the model

### 6.3 Active Learning Experiments

We now describe our experiments with the active learning criteria. In order to test the results of active learning on linear regression, we used the set  $\mathcal{T}$  as the test dataset as in the previous case. Initially, only 10 instances from  $\mathcal{F}$  labeled by all annotators were used as the training set  $\mathcal{D}$ .  $\mathcal{F} \setminus \mathcal{D}$  was used as the unlabeled set  $\mathcal{U}$ . At every step of active learning, the label  $y_{kj^*}$  of one instance  $\mathbf{x}_k$  was procured from an annotator  $j^*$ , chosen using Algorithm 1. The model was relearned using the new training set  $\mathcal{D} = \mathcal{D} \cup \{(\mathbf{x}_k, y_{kj^*})\}$ . The RMSE was calculated on  $\mathcal{T}$  and the results were plotted at every step. We also plotted the regret for Algorithm 1 at every step. The experiments were repeated for 10 different splits of the dataset. The test RMSE when the set  $\mathcal{F}$  was used for training (so that  $\mathcal{D} = \mathcal{F}$ ) was also plotted. This error is the best achievable error in the crowdsourcing scenario.

To the best of our knowledge, our work is the first attempt towards active learning for regression from the crowd and therefore there are no other baselines in the literature to compare our method against. However we have used the following baselines for comparison:

- (1)Random: Random selection of instances and annotators.
- (2)Instance: Algorithm 1 for selecting the instances and random selection of annotators.

(3)Single Source AL: The labels were provided by a single source with negligible noise. Active selection of instances was performed using uncertainty sampling. The RMSE and regret plots are provided in Figure 3. Clearly the Robust UCB strategy outperforms ‘Random’ as well as ‘Instance’ with respect to RMSE as well as regret and approaches the ‘Single Source AL’ with fewer number of labeled examples.

*Remark 3.* Our active learning algorithm demonstrates a superior performance with just a few additional labels (Figure 3). A similar trend was observed for the rest of the curve, which was omitted in the plots for the sake of clarity.

## 7 Conclusions and Future Work

We set up a Bayesian framework to infer the parameters of linear regression using crowds. As closed form Bayesian solution is intractable, we used approximation schemes. To improve this initially learnt regression model, we used various active learning techniques and studied their theoretical foundations. We established the connections with MAB algorithms and explored the use of Robust UCB for annotator selection in active learning, providing theoretical guarantees and also performing a wastage analysis. Next, we introduced a payment scheme for annotators to ensure that they put in their best efforts while labeling the data. Our experiments on real data show the efficacy of our techniques.

Our approach of Bayesian learning, MAB algorithm for annotator selection, uncertainty sampling for instance selection and design of quality compatible mechanisms to elicit best efforts from crowd workers is applicable for a wide range of tasks like classification, ordinal regression etc. It would be interesting to study the suitability of various MAB algorithms depending on the form of the distributions used to model the annotators’ qualities. Modeling the subjectivity of the annotators, their dynamic entry and exit, and the design of incentives in these scenarios is also challenging.

## Bibliography

- [1] I. Abraham, O. Alonso, V. Kandylas, and A. Slivkins. Adaptive crowdsourcing algorithms for the bandit survey problem. In *COLT*, pages 882–910, 2013.
- [2] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *JMLR*, 47(2-3):235–256, 2002.
- [3] M. J. Beal. *Variational Algorithms for Approximate Bayesian Inference*. PhD thesis, Gatsby Computational Neuroscience Unit, University College London, 2003.
- [4] W. Bi, L. Wang, J. T. Kwok, and Z. Tu. Learning to predict from crowd-sourced data. In *UAI*, 2014.
- [5] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [6] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., 2006.
- [7] S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic and non-stochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122, 2012.
- [8] S. Bubeck, N. Cesa-Bianchi, and G. Lugosi. Bandits with heavy tail. *IEEE Transactions on Information Theory*, 59(11):7711–7717, 2013.
- [9] R. Burbidge, J. J. Rowland, and R. D. King. Active learning for regression based on query by committee. In *IDEAL*, pages 209–218, 2007.
- [10] W. Cai, Y. Zhang, and J. Zhou. Maximizing expected model change for active learning in regression. In *ICDE*, pages 51–60, 2013.
- [11] Y. Cai, C. Daskalakis, and C. H. Papadimitriou. Optimum statistical estimation with strategic data sources. In *COLT*, pages 280–296, 2015.
- [12] D. A. Cohn, Z. Ghahramani, and M. I. Jordan. Active learning with statistical models. *JAIR*, 4(1):129–145, 1996.
- [13] P. Dayama, B. Narayanaswamy, D. Garg, and Y. Narahari. Truthful interval cover mechanisms for crowdsourcing applications. In *AAMAS*, pages 1091–1099, 2015.
- [14] O. Dekel, F. Fischer, and A. D. Procaccia. Incentive compatible regression learning. *Journal of Computer and System Sciences*, 76(8):759–777, 2010.
- [15] O. Dekel, C. Gentile, and K. Sridharan. Selective sampling and active learning from single and multiple teachers. *JMLR*, 13:2655–2697, 2012.
- [16] V. Fedorov. *Theory of Optimal Experiments*. Academic Press, New York, 1972.
- [17] P. Groot, A. Birlutiu, and T. Heskes. Learning from multiple annotators with gaussian processes. In *ICANN*, pages 159–164, 2011.
- [18] V. Grover. Active learning and its application to heteroscedastic problems. Master’s thesis, University of Alberta, 2009.
- [19] M. Lichman. UCI machine learning repository, 2013. URL <http://archive.ics.uci.edu/ml>.

- [20] D. V. Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4), 1956.
- [21] D. J. MacKay. Information-based objective functions for active data selection. *Neural Computation*, 4(4):590–604, 1992.
- [22] J. Neufeld, A. György, C. Szepesvári, and D. Schuurmans. Adaptive monte carlo via bandit allocation. In *ICML*, pages 1944–1952, 2014.
- [23] V. Raykar and P. Agrawal. Sequential crowdsourced labeling as an epsilon-greedy exploration in a markov decision process. In *AISTATS*, pages 832–840, 2014.
- [24] V. Raykar, S. Yu, L. H. Zhao, A. Jerebko, C. Florin, G. H. Valadez, L. Bogoni, and L. Moy. Supervised learning from multiple experts: Whom to trust when everyone lies a bit. In *ICML*, pages 889–896, 2009.
- [25] V. Raykar, S. Yu, L. H. Zhao, G. H. Valadez, C. Florin, L. Bogoni, and L. Moy. Learning from crowds. *JMLR*, 11:1297–1322, 2012.
- [26] K. Ristovski, D. Das, V. Ouzienko, Y. Guo, and Z. Obradovic. Regression learning with multiple noisy oracles. In *ECAI*, pages 445–450, 2010.
- [27] F. Rodrigues, F. C. Pereira, and B. Ribeiro. Gaussian process classification and active learning with multiple annotators. In *ICML*, pages 433–441. 2014.
- [28] J. Roeder, B. Nadler, K. Kunzmann, and F. Hamprecht. Active learning with distributional estimates. In *UAI*, pages 715–725, 2012.
- [29] N. Roy and A. McCallum. Toward optimal active learning through sampling estimation of error reduction. In *ICML*, pages 441–448, 2001.
- [30] S. Sen, A. Ridgway, and M. Ripley. Adaptive budgeted bandit algorithms for trust development in a supply-chain. In *AAMAS*, pages 137–144, 2015.
- [31] B. Settles. Active learning literature survey. Technical report, University of Wisconsin–Madison, 2010.
- [32] B. Settles, M. Craven, and S. Ray. Multiple-instance active learning. In *NIPS*, pages 1289–1296. 2008.
- [33] H. S. Seung, M. Opper, and H. Sompolinsky. Query by committee. In *COLT*, pages 287–294, 1992.
- [34] L. Tran-Thanh, T. D. Huynh, A. Rosenfeld, S. D. Ramchurn, and N. R. Jennings. Budgetfix: Budget limited crowdsourcing for interdependent task allocation with quality guarantees. In *AAMAS*, pages 901–908, 2014.
- [35] F. L. Wauthier and M. I. Jordan. Bayesian bias mitigation for crowdsourcing. In *NIPS*, pages 1800–1808. 2011.
- [36] P. Zhao, S. Hoi, and J. Zhuang. Active learning with expert advice. In *UAI*, 2013.